

# Método de Newton

Marina Andretta

ICMC-USP

12 de agosto de 2014

Baseado no livro Numerical Optimization, de J. Nocedal e S. J. Wright.

O **método de Newton** para resolução de problemas de minimização irrestrita é um método de busca linear que usa a direção

$$p_k = -(\nabla^2 f_k)^{-1} \nabla f_k.$$

(1)

Note que, neste caso,

$$p_k^T \nabla f_k = -\nabla f_k^T (\nabla^2 f_k)^{-1} \nabla f_k.$$

Ou seja, se a Hessiana  $\nabla^2 f_k$  for definida positiva,  $p_k$  é uma direção de descida.

Além disso, se as matrizes  $\nabla^2 f_k$  têm número de condição uniformemente limitado, ou seja, existe uma constante  $M$  tal que

$$\|\nabla^2 f_k\| \|\nabla^2 f_k^{-1}\| \leq M, \quad \forall k,$$

vale que

$$\cos \theta_k \geq 1/M.$$

Usando a condição de Zoutendijk, temos que

$$\lim_{k \rightarrow \infty} \|\nabla f_k\| = 0.$$

Portanto, o método de Newton produz uma sequência de gradientes que converge para zero, dado que as Hessianas  $\nabla^2 f_k$  sejam definidas positivas com número de condição limitado uniformemente e que o método use uma busca linear que satisfaça as condições de Wolfe.

Pode-se mostrar também que a sequência gerada pelo método de Newton converge quando o tamanho de passo é calculado usando *backtracking* e satisfaz a condição de Armijo.

# Ordem de convergência (local)

Como a matriz Hessiana  $\nabla^2 f_k$  pode não ser definida positiva, a direção  $p_k$  pode não ser de descida. Futuramente serão discutidas técnicas para contornar este problema.

Vejam agora apenas **propriedades da ordem de convergência local do método de Newton**. Sabemos que, para todo ponto  $x$  numa vizinhança da solução  $x^*$  que tem  $\nabla^2 f(x^*)$  definida positiva,  $\nabla^2 f(x)$  também é definida positiva.

Para esta vizinhança, o **método de Newton está bem definido e converge quadraticamente** dado que os tamanhos de passo  $\alpha_k$  sejam, a partir de uma certa iteração, sempre 1.

**Teorema 1:** *Suponha que  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  tenha segunda derivada contínua e que a Hessiana  $\nabla^2 f(x)$  seja Lipschitz contínua em uma vizinhança da solução  $x^*$ , na qual as condições suficientes de segunda ordem são satisfeitas. Considere a iteração da forma  $x_{k+1} = x_k + p_k$  com  $p_k = -\nabla^2 f_k^{-1} \nabla f_k$ .*

*Então,*

- *se o ponto inicial  $x_0$  está suficientemente próximo de  $x^*$ , a sequência de iterandos converge a  $x^*$ ;*
- *a ordem de convergência de  $\{x_k\}$  é quadrática; e*
- *a sequência da norma dos gradientes  $\{\|\nabla f_k\|\}$  converge quadraticamente para 0.*



# Ordem de convergência (local)

Se o método usa uma busca linear que busca satisfazer as condições de Wolfe e sempre tenta o tamanho unitário, o teorema a seguir garante que este tamanho de passo sempre será aceito a partir de uma certa iteração (já que o limite (2) sempre vale quando a direção  $p_k$  é calculada pelo método de Newton).

**Teorema 2:** *Suponha que  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  tenha terceira derivada contínua. Considere uma iteração da forma  $x_{k+1} = x_k + \alpha_k p_k$ , onde  $p_k$  é uma direção de descida e  $\alpha_k$  satisfaz as condições de Wolfe com  $c_1 \leq 0.5$ . Se a sequência  $\{x_k\}$  converge a um ponto  $x^*$  tal que  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é definida positiva e se a direção de busca satisfaz*

$$\lim_{k \rightarrow \infty} \frac{\|\nabla f_k + \nabla^2 f_k p_k\|}{\|p_k\|} = 0, \quad (2)$$

então:

- 1 o tamanho de passo  $\alpha_k = 1$  é admissível para todo  $k$  maior do que um certo índice  $k_0$ ; e
- 2 se  $\alpha_k = 1$  para todo  $k > k_0$ ,  $\{x_k\}$  converge a  $x^*$  superlinearmente.